DEPARTMENT OF PHILOSOPHY TALK

TIFFANY ZHU

"HOW SHOULD AI TALK ABOUT US?"

GRADUATE STUDENT AT THE UCI DEPARTMENT OF PHILOSOPHY

Generative artificial intelligence technologies are technologies of abstraction and prediction. When image generators are asked to depict groups like women and immigrants, the results almost invariably contain stereotypes. Language models face similar challenges when responding to inquiries about people based on social positions and group affiliations. Can genAI communicate truth-sensitive and justice-conducive information about groups of people in a world patterned by categorization and discrimination?

In this talk, I address this question by evaluating how ChatGPT responds to prompts containing generic generalizations such as "Do immigrants hold lower-paying jobs?" My analysis reveals troubling responses, including over-censorship and the minimization of structural inequities. I then sketch and evaluate five potential solutions for mitigating human-AI miscommunication and harmful social consequences. Overall, I advance three claims regarding the ethics of generative artificial intelligence. First, they are not set up to portray people and social reality in justice-conducive ways. Second, preventing or correcting flawed portrayals requires complex solutions. Finally, genAI technologies are not human, and that can be a good thing.

JANUARY 27TH, 2025 2-4PM HIB 55

UCI Department of Philosophy

Please email philos@uci.edu for further questions